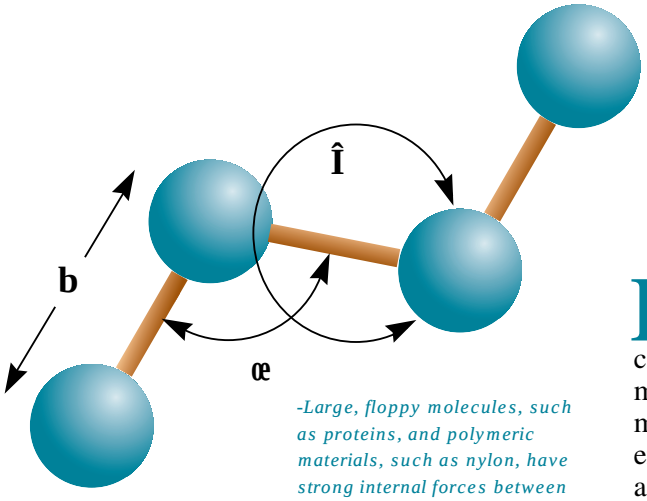
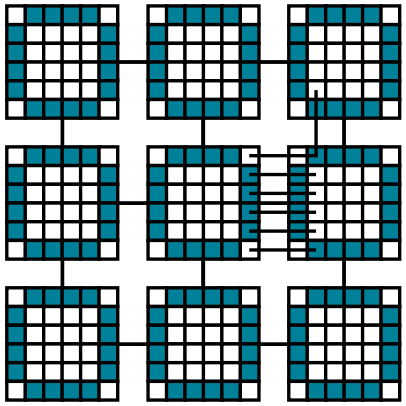
Modeling Large Molecular Systems

Laboratory scientists are developing a set of computational tools to simulate large, floppy molecules on the new Cray T3D parallel supercomputer to extend the capabilities of DuPont and Bristol-Myers Squibb.



-Large, floppy molecules, such as proteins, and polymeric materials, such as nylon, have strong internal forces between atoms that are covalently bonded and very long-range forces with atoms that are far away.



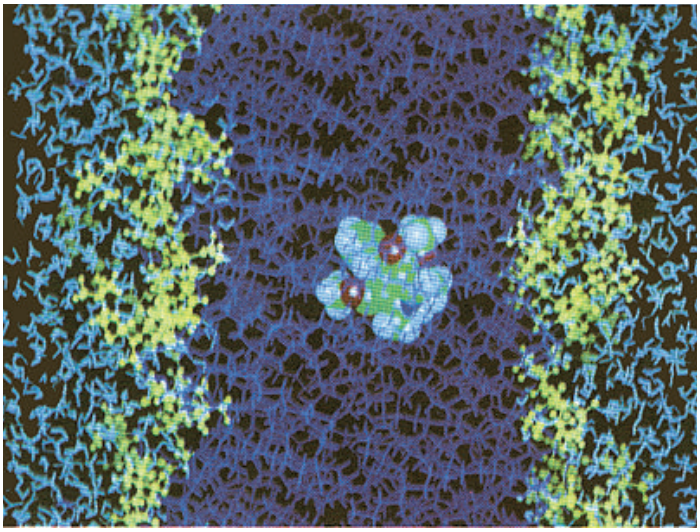
We map large molecular dynamics problems to a parallel computer by dividing the problem into smaller regions or domains.

BOTH the DOE and U.S. chemical and pharmaceutical industries need to accelerate the rate at which they can create new organic-based materials. Examples of the materials being studied include biological molecules and membranes, liquid crystals, optical thin films, and energetic materials. Conventional supercomputers now assist materials designers by giving them the ability to routinely model tens of thousands of atoms for hundreds of picoseconds (1 ps = 10^{-12} s).

However, to accurately simulate many important biological and physical phenomena, we require even more computational power. To simulate large molecular systems with long-range forces for periods as long as tens of nanoseconds and beyond (1 ns = 10^{-9} s), we need at least one or two orders of magnitude more computational power. This power will become available with the next generation of supercomputer, which uses many microprocessors to simultaneously solve a single problem.

Earlier this year, the Laboratory signed a CRADA with three industrial partners and Sandia National Laboratories/New Mexico to address specific computational issues. The goal of the CRADA is to develop the software needed for the newest massively parallel supercomputers in studies of large molecular systems with long-range forces.

In addition to the two national laboratories, the three-year program also involves DuPont, Bristol-Myers Squibb, and Cray Research. DuPont, based in Wilmington, Delaware, has had an ongoing effort in computational chemistry since the mid-1970s. This research and technology-based global company offers advanced technologies and products, such as petroleum, chemicals, polymers, and fibers. Bristol-Myers Squibb is one of the world's largest pharmaceutical companies. Its Pharmaceutical Research Institute, based in Princeton,



A simulation, performed at the Bristol-Myers Squibb Pharmaceutical Research Institute in Princeton, N.J., of a drug molecule at the center of a biomembrane.[1]

New Jersey, has a large computational chemistry effort that assists in understanding biochemical phenomena and in the development of new pharmaceuticals. Cray Research Inc., based in Eagan, Minnesota, is the nation's premiere supercomputer hardware manufacturer. The agreement with these partners will allow access to expertise and computing power that are beyond the scope of any single company or laboratory.

The use of computational science to understand, model, and develop new materials has a proven track record of getting products to market faster. Nevertheless, industry has a pressing need to expand the role that computational science plays. Thus, the partners in the new five-way CRADA will develop a new set of computational tools to examine the behavior of new materials using massively parallel supercomputers. One such computer is the T3D being developed by Cray Research. (The DOE has now leased to LLNL a T3D machine with 128 processors as part of the Industrial Computing Initiative.)

An example of the type of problem we are addressing is the transport of a drug molecule through a biomembrane. In essence, a large molecular dynamics problem such as this is mapped to a parallel computer by dividing the problem into smaller regions or domains. (See the article on p. 13 for an introduction to molecular dynamics modeling.) Each domain, containing many atoms, is treated as a separate molecular dynamics problem and is assigned to a separate processor. That is, a processor is responsible for evolving all the atoms within its domain. At the boundary of each domain are atoms that belong to neighboring domains and different processors. Care must be taken to ensure that the boundary information is communicated at each time step in the calculation so that the correct forces on the atoms are applied.

Another difficulty that arises for large biological systems with water present is that the interaction forces between atoms are very long range. In effect, every atom interacts with every other atom. This situation places serious limitations on the number of atoms we can simulate. To solve the problem, laboratory scientists are exploring an alternative method known as the fast multipole method. The main idea is that, instead of calculating the interaction of every atom that is far away, the code calculates the interaction of an atom with the field produced by distant atoms. The method saves considerable computer time for very large systems. Such savings are crucial to extend our present capabilities to solve key problems facing U.S. chemical and pharmaceutical industries.

Reference

1. T. R. Stouch, H. E. Alper, and D. Bassolino-Klimas, "Supercomputing Studies of Biomembranes," *Supercomputer Applications and High Performance Computing* **8**, 6-23 (1994).

For further information contact James Belak (510) 422-6061 or Roy Pollock (510) 422-4088.